

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model1

Run on: June 2, 2004, 11:28:57 ; Search time 6.03261 Seconds
(without alignments)
319.364 Million cell updates/sec

Title: US-09-657-336A-179
Perfect score: 181
Sequence: 1 NLIRALIAQOHLLQLTWQIKQLQARILAVERYLKQD 37

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 90 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	174	96.1	853	1 ENV_HV1EL	P04581 human immun
2	174	96.1	853	1 ENV_HV1MF	P19551 human immun
3	174	96.1	853	1 ENV_HV122	P12457 human immun
4	174	96.1	855	1 ENV_HV126	P04580 human immun
5	174	96.1	856	1 ENV_HV1B1	P03375 human immun
6	174	96.1	856	1 ENV_HV1H2	P04578 human immun
7	174	96.1	856	1 ENV_HV1H3	P04624 human immun
8	174	96.1	856	1 ENV_HV1LV	P070626 human immun
9	174	96.1	856	1 ENV_HV1PV	P03376 human immun
10	174	96.1	861	1 ENV_HV1BR	P03377 human immun
11	173	95.6	847	1 ENV_HV1S1	P19550 human immun
12	173	95.6	855	1 ENV_HV1OY	P20888 human immun
13	171	94.5	848	1 ENV_HV1JR	P20871 human immun
14	171	94.5	856	1 ENV_HV1AN	P05877 human immun
15	170	93.9	843	1 ENV_HV1Y2	P35961 human immun
16	170	93.9	846	1 ENV_HV1ND	P18799 human immun
17	170	93.9	851	1 ENV_HV1B8	P04582 human immun
18	170	93.9	852	1 ENV_HV1S3	P19549 human immun
19	170	93.9	855	1 ENV_HV1A2	P03378 human immun
20	170	93.9	856	1 ENV_HV1SC	P05878 human immun
21	170	93.9	856	1 ENV_HV1W1	P31872 human immun
22	170	93.9	856	1 ENV_HV1ZH	P05881 human immun
23	170	93.9	861	1 ENV_HV1KB	P31819 human immun
24	170	93.9	865	1 ENV_HV1RH	P04579 human immun
25	170	93.9	867	1 ENV_HV1J3	P12489 human immun
26	170	93.9	868	1 ENV_HV1C4	P05879 human immun
27	169	93.4	859	1 ENV_HV1A4	P04583 human immun
28	167	92.3	847	1 ENV_HV1W2	P05880 human immun
29	165	91.2	863	1 ENV_HV1B8	P05880 human immun
30	162	89.5	852	1 ENV_HV1BN	P12488 human immun
31	159	87.8	854	1 ENV_SIVC2	P12781 chimpanzee
32	125	69.1	865	1 ENV_SIVAT	P05886 simian immu
33	125	69.1	877	1 ENV_SIVAG	P27977 simian immu

34	122	67.4	854	1 ENV_SIVAI	Q02837 simian immu
35	116	62.1	768	1 ENV_SIVAI	P27757 simian immu
36	113	62.4	712	1 ENV_HV2S2	P32536 human immu
37	113	62.4	851	1 ENV_HV2D1	P17755 human immu
38	113	62.4	851	1 ENV_HV2G1	P18040 human immu
39	113	62.4	856	1 ENV_HV2NZ	P05883 human immu
40	113	62.4	859	1 ENV_HV2ST	P20872 human immu
41	112	61.9	859	1 ENV_HV2CA	P24105 human immu
42	111	61.3	380	1 ENV_SIVM2	P08810 simian immu
43	111	61.3	858	1 ENV_HV2RO	P04570 simian immu
44	111	61.3	885	1 ENV_SIVS4	P12492 simian immu
45	111	61.3	889	1 ENV_SIVSP	P19503 simian immu
46	110	60.8	882	1 ENV_SIVM1	P05885 simian immu
47	109	60.2	859	1 ENV_HV2D2	P15831 human immu
48	108	59.7	880	1 ENV_SIVML	P11267 simian immu
49	107	59.1	821	1 ENV_SIVGB	P22380 simian immu
50	107	59.1	846	1 ENV_HV2SB	P12449 human immu
51	107	59.1	881	1 ENV_SIVMK	P05884 simian immu
52	106	58.6	857	1 ENV_HV2KR	Q74126 human immu
53	106	58.6	860	1 ENV_HV2BE	P18094 human immu
54	52	28.7	2564	1 SPCQ_HUMAN	Q91254 human sapien
55	51.5	28.5	924	1 HXK3_RAT	P27926 rattus norv
56	50.5	27.9	445	1 EX7L_STRAM	Q95X00 staphylococ
57	50.5	27.9	1411	1 EBAL_HUMAN	Q15075 homo sapien
58	50	27.6	1938	1 MYR4_RABIT	Q28641 oryctolagus
59	49.5	27.3	581	1 FRIZ_DROME	P18537 drosophila
60	49.5	27.3	583	1 FRIZ_DROVI	Q24760 drosophila
61	49	27.1	236	1 GT6_SCHMA	P64335 schistosoma
62	49	27.1	1756	1 PEP1_HUMAN	O60437 homo sapien
63	49	27.1	1938	1 MY5_AEGIR	P24733 aequipecten
64	49	27.1	1939	1 MYR4_HUMAN	Q9Y623 homo sapien
65	49	27.1	4349	1 DYHC_FUSO	P78716 fusarium so
66	49	27.1	8797	1 SNEI_HUMAN	O8N691 homo sapien
67	48.5	26.8	790	1 EBAL_MOUSE	O8b166 mus musculu
68	48	26.5	213	1 Y914_THEMA	Q9X016 thermocoga
69	48	26.5	305	1 OTC_STMP7	Q935Y4 synechococ
70	48	26.5	314	1 SEH1_HUMAN	Q9H418 homo sapien
71	48	26.5	551	1 YD25_YEAST	Q07657 saccharomyc
72	48	26.5	702	1 AT11_VAVR	P34011 variola vir
73	48	26.5	906	1 CTN1_MOUSE	P35221 homo sapien
74	48	26.5	906	1 CTN1_MOUSE	P26231 mus musculu
75	48	26.5	1379	1 M3K5_MOUSE	O35099 mus musculu
76	47.5	26.2	253	1 CYSH_VIBCH	Q9KXJ2 vibrio chol
77	47.5	26.2	962	1 ARVC_HUMAN	O00192 homo sapien
78	47.5	26.2	969	1 ARVC_MOUSE	P98203 mus musculu
79	47.5	26.2	1319	1 S0S1_MOUSE	O62245 mus musculu
80	47.5	26.2	1333	1 S0S1_HUMAN	O07889 homo sapien
81	47.5	26.2	1955	1 PUMA_PARKN	O61308 parascaris
82	47	26.0	376	1 O43A_DROME	P81917 drosophila
83	47	26.0	380	1 OXAA_MYCLE	Q50205 mycobacteri
84	47	26.0	418	1 DADI_HILLO	O98508 rhizobium 1
85	47	26.0	851	1 STA2_HUMAN	P52630 homo sapien
86	47	26.0	1935	1 MYSS_CYPCA	O90339 cyprinus ca
87	46.5	25.7	132	1 RS8_CIOAB	Q976J2 clostridium
88	46.5	25.7	134	1 Y652_METUA	Q58068 metheriococ
89	46.5	25.7	467	1 EUTE_ECOLI	P77445 escherichia
90	46.5	25.7	467	1 EUTE_SALTY	P41793 salmonella

ALIGNMENTS

RESULT 1
ENV_HV1EL STANDARD; PRT; 853 AA.
AC P04581;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BLI isolate) (HIV-1).

OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=11689;
RN [1]
RP MEDLINE=66245056; PubMed=2424612;
RX Alizon M., Main-Hosson S., Montagnier L., Sonigo P.;
RA "Genetic variability of the AIDS virus: nucleotide sequence analysis
RT of two isolates from African patients";
RL Cell 46:63-74(1986).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K03454; AAA44329.1; -;
DR EMBL; A07108; CAA00616.1; -;
DR HIV; K03454; ENVSELI.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
KW Signal.
FT CHAIN 1 31 BY SIMILARITY.
FT SIGNAL 32 508 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 509 853 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 206 BY SIMILARITY.
FT DISULFID 125 197 BY SIMILARITY.
FT DISULFID 130 154 BY SIMILARITY.
FT DISULFID 219 248 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 297 330 BY SIMILARITY.
FT DISULFID 376 442 BY SIMILARITY.
FT DISULFID 383 416 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 137 137 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 143 143 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 153 153 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 157 157 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 183 183 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 188 188 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 353 353 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 394 394 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 411 411 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 462 462 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 608 608 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 613 613 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 622 622 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 634 634 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 853 AA; 96721 MW; F9C864DAAD07A5 CRC64;

Best Local Similarity 97.3%; Pred. No. 2, 4e-16;
Matches 36; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 NLRRAEAOQHLLQLTWMOIKOLQARILAVERYLKQD 37
Db 551 NLRRAEAOQHLLQLTWMOIKOLQARILAVERYLKQD 587
RESULT 2
ENV_HVIMF STANDARD; PRT; 853 AA.
ID ENV_HVIMF
AC P19551;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP120 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (MVA isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=11704;
RN [1]
RP MEDLINE=90317877; PubMed=1695254;
RX Stevenson M., Haggerty S., Lamonica C., Mann A.M., Meier C.,
RA Maslak A.;
RT "Cloning and characterization of human immunodeficiency virus type 1
RT variants diminished in the ability to induce syncytium-independent
RT cytolysis";
RL J. Virol. 64:3792-3803(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M3943; AAA44850.1; -;
DR PDB; 1AIX; 16-JUN-97.
DR HIV; M3943; ENVSEMA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 30
FT CHAIN 31 509 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 510 853 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 54 74 BY SIMILARITY.
FT DISULFID 119 203 BY SIMILARITY.
FT DISULFID 126 194 BY SIMILARITY.
FT DISULFID 131 157 BY SIMILARITY.
FT DISULFID 216 245 BY SIMILARITY.
FT DISULFID 226 237 BY SIMILARITY.
FT DISULFID 294 329 BY SIMILARITY.
FT DISULFID 376 443 BY SIMILARITY.
FT DISULFID 383 416 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).

Query Match

96.1%; Score 174; DB 1; Length 853;

```

FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 404 404 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 446 446 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 461 461 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 609 609 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 614 614 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 635 635 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 672 672 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 748 748 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 814 814 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 853 AA; 96912 MW; 3377893B6F22A8A CRC64;

Query Match 96.1%; Score 174; DB 1; Length 853;
Best Local Similarity 97.3%; Pred. No. 2,4e-16;
Matches 36; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLRRAIEAOQHLLQLTWQIKQARILAVERYLKQ 37
Db 552 NLRRAIEAOQHLLQLTWQIKQARILAVERYLKQ 588

RESULT 3
ENV_HV126 STANDARD; PRT; 853 AA.
ID ENV_HV126
AC P12487;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Envelope glycoprotein GP160 precursor [contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (Z2/CDC-234 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentiviruses.
OX NCBI_TaxID=11683;
RN [1]
RP SEQUENCE FROM N.A.
RA Theodore T., Buckler-White A.;
RL Submitted (NOV-1988) to the HIV data bank.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M22639; AAA45370.1; -.
DR PIR; S54384; S54384.
DR HIV; M22639; ENV52226.
DR InterPro: IPR000328; Env GP41.
DR InterPro: IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KM Signal.
FT SIGNAL 1 31 BY SIMILARITY.
FT CHAIN 32 508 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 509 853 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 206 BY SIMILARITY.
FT DISULFID 125 197 BY SIMILARITY.
FT DISULFID 130 154 BY SIMILARITY.
FT DISULFID 219 248 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 297 330 BY SIMILARITY.
FT DISULFID 376 442 BY SIMILARITY.

```

```

FT DISULFID 383 415 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 137 137 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 338 338 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 353 353 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 462 462 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 608 608 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 613 613 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 622 622 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 634 634 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 853 AA; 97043 MW; 849B08BCBAFF7008 CRC64;

Query Match 96.1%; Score 174; DB 1; Length 853;
Best Local Similarity 97.3%; Pred. No. 2,4e-16;
Matches 36; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLRRAIEAOQHLLQLTWQIKQARILAVERYLKQ 37
Db 551 NLRRAIEAOQHLLQLTWQIKQARILAVERYLKQ 587

RESULT 4
ENV_HV126 STANDARD; PRT; 855 AA.
ID ENV_HV126
AC P04580;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (Zaire 6 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentiviruses.
OX NCBI_TaxID=11708;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87248097; PubMed=3036660;
RA Strinivasan A., Anand R., York D., Ranganathan P., Feorino P.,
RA Sanchez-Pescador R.,
RA Sanchez-Pescador R.,
RA Kalyanaraman V.S., Luciw P.A.,
RT "Molecular characterization of human immunodeficiency virus from
RT Zaire; nucleotide sequence analysis identifies conserved and variable
RT domains in the envelope gene.";
RT Gene 52:71-82(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

[illegible]

DE				glycoprotein (GP120); Transmembrane glycoprotein (GP41)).
RN	ENV.			.
OC	Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).			
CC	Viruses; Retroid viruses; Retroviridae; Lentivirs.			
OX	NCB1_TaxID=11678;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=9511123; Pubmed=2576615;			
RA	Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R., Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A., Baumeister K.K., Ivanoff L., Petteway S.R. Jr., Pearson M.L., Latemaeijer J.C., Papas T.S., Graybe V., Chang N.T., Gallo R.C., Wong-Staal F.; "Complete nucleotide sequence of the AIDS virus, HTLV-II." Nature 313:277-284(1985). [2] DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.			
RX	MEDLINE=9028515; PubMed=235506;			
RA	Leonard C.K., Spellman M.W., Riddle L., Harris R.J., Thomas J.N., Gregory T.J.; "Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type I recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells."; J. Biol. Chem. 265:10373-10382(1990).			
RL	J. Biol. Chem. 265:10373-10382(1990).			
RT	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - CC between the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).			
CC	or send an email to license@isb-sib.ch).			
CC	EMBL; M15654; AAAA4205.1; -. PIR; A03973; VCLJH3.			
DR	HIV; M15654; ENVSBHI02.			
DR	IInterPro; IPRO00328; Env GP41.			
DR	IInterPro; IPRO00777; GP120.			
DR	Pfam; PF00516; GP120; 1.			
DR	PFam; PF00517; GP41; 1.			
KW	AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal.			
FT	SIGNAL.	30		
FT	CHAIN	31	511	
FT	CHAIN	512	856	EXTERIOR MEMBRANE GLYCOPROTEIN. TRANSMEMBRANE GLYCOPROTEIN.
FT	DISULFID	54	74	
FT	DISULFID	119	205	
FT	DISULFID	126	196	
FT	DISULFID	131	157	
FT	DISULFID	218	247	
FT	DISULFID	228	239	
FT	DISULFID	296	331	
FT	DISULFID	378	445	
FT	DISULFID	385	418	
FT	CARBOHYD	88	88	N-LINKED (GLCNAC . .)
FT	CARBOHYD	136	136	N-LINKED (GLCNAC . .)
FT	CARBOHYD	141	141	N-LINKED (GLCNAC . .)
FT	CARBOHYD	156	156	N-LINKED (GLCNAC . .)
FT	CARBOHYD	160	160	N-LINKED (GLCNAC . .)
FT	CARBOHYD	186	186	N-LINKED (GLCNAC . .)
FT	CARBOHYD	197	197	N-LINKED (GLCNAC . .)
FT	CARBOHYD	230	230	N-LINKED (GLCNAC . .)
FT	CARBOHYD	234	234	N-LINKED (GLCNAC . .)
FT	CARBOHYD	241	241	N-LINKED (GLCNAC . .)
FT	CARBOHYD	262	262	N-LINKED (GLCNAC . .)
FT	CARBOHYD	276	276	N-LINKED (GLCNAC . .)
FT	CARBOHYD	289	289	N-LINKED (GLCNAC . .)
FT	CARBOHYD	295	295	N-LINKED (GLCNAC . .)
FT	CARBOHYD	301	301	N-LINKED (GLCNAC . .)
FT	CARBOHYD	332	332	N-LINKED (GLCNAC . .)
FT	CARBOHYD	339	339	N-LINKED (GLCNAC . .)
FT	CARBOHYD	356	356	N-LINKED (GLCNAC . .)

```

FT CARBOHYD 386 386 N-LINKED (GLCNAC . . .)
FT CARBOHYD 392 392 N-LINKED (GLCNAC . . .)
FT CARBOHYD 397 397 N-LINKED (GLCNAC . . .)
FT CARBOHYD 406 406 N-LINKED (GLCNAC . . .)
FT CARBOHYD 448 448 N-LINKED (GLCNAC . . .)
FT CARBOHYD 463 463 N-LINKED (GLCNAC . . .)
FT CARBOHYD 611 611 N-LINKED (GLCNAC . . .)
FT CARBOHYD 616 616 N-LINKED (GLCNAC . . .)
FT CARBOHYD 625 625 N-LINKED (GLCNAC . . .)
FT CARBOHYD 637 637 N-LINKED (GLCNAC . . .)
FT CARBOHYD 674 674 N-LINKED (GLCNAC . . .)
FT CARBOHYD 750 750 N-LINKED (GLCNAC . . .)
FT CARBOHYD 816 816 N-LINKED (GLCNAC . . .)
SQ SEQUENCE 856 AA; 97224 MW; 0BFFB1A18931BB27 CRC64;

Query Match 96.1%; Score 174; DB 1; Length 856;
Best Local Similarity 97.3%; Pred. No. 2,4e-16;
Matches 36; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Dy 1 NLRAIEAQOHLLQITWQIQOLQRIIAVERYLKDDQ 37
Db 554 NLRAIEAQOHLLQITWQIQOLQRIIAVERYLKDDQ 590

RESULT 6
ENV_HV1H2 STANDARD; PRT; 856 AA.
ID ENV_HV1H2
AC P04578; O09779;
DT 13-AUG-1987 (Rel. 05, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
DE ENV.
OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11706;
RX MEDLINE=87299196; PubMed=3040055;
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RT "Complete nucleotide sequences of functional clones of the AIDS
RT virus."
RL AIDS Res. Hum. Retroviruses 3:57-69(1987).
RL [2]
RA REVISIONS.
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K03455; AAB50262.1; -
DR EMBL; AF038399; AAB99976.1; -
DR EMBL; AF033819; AAC82596.1; -
DR PDB; 1DP4; 26-JAN-00.
DR PDB; 1DE5; 26-JAN-00.
DR PDB; 1DIB; 02-APR-00.
DR PDB; 1G9M; 27-DEC-00.
DR PDB; 1GCL; 19-AUG-98.
DR PDB; 1GZL; 10-OCT-02.
DR PDB; 1K33; 10-OCT-01.
DR PDB; 1K34; 10-OCT-01.
DR HIV; K03455; ENVSHXB2.
DR InterPro; IPR000328; ENV_GP41.
DR InterPro; IPR000777; GP120.

```

Accession	Protein	Signal	Score	Length	DB	Query Match	Best Local Similarity	Matches	Conservative	Mismatches	Indels	Gaps
DR PFam: PF00516; GP120; 1.	Human immunodeficiency virus type 1 (HXB3 isolate) (HIV-1).	Signal	30	856	DB 1;	96.1%;	Score 174;	36;	97.3%;	0;	0;	0;
DR PFam: PF00517; GP41; 1.	Human immunodeficiency virus type 1 (HXB3 isolate) (HIV-1).	Signal	30	856	DB 1;	96.1%;	Score 174;	36;	97.3%;	0;	0;	0;
KM AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;												
KM 3D-structure.												
FT CHAIN 1	EXTERIOR MEMBRANE GLYCOPROTEIN.		31	511								
FT CHAIN 2	TRANSMEMBRANE GLYCOPROTEIN.		512	856								
FT DISULFID 54	BY SIMILARITY.		54	74								
FT DISULFID 119	BY SIMILARITY.		119	205								
FT DISULFID 126	BY SIMILARITY.		126	196								
FT DISULFID 131	BY SIMILARITY.		131	157								
FT DISULFID 218	BY SIMILARITY.		218	247								
FT DISULFID 228	BY SIMILARITY.		228	239								
FT DISULFID 296	BY SIMILARITY.		296	331								
FT DISULFID 378	BY SIMILARITY.		378	445								
FT DISULFID 385	BY SIMILARITY.		385	418								
FT CARBOHYD 88	N-LINKED (GLCNAC. . .)	(POTENTIAL).	88	136								
FT CARBOHYD 136	N-LINKED (GLCNAC. . .)	(POTENTIAL).	136	141								
FT CARBOHYD 141	N-LINKED (GLCNAC. . .)	(POTENTIAL).	141	156								
FT CARBOHYD 156	N-LINKED (GLCNAC. . .)	(POTENTIAL).	156	160								
FT CARBOHYD 160	N-LINKED (GLCNAC. . .)	(POTENTIAL).	160	186								
FT CARBOHYD 186	N-LINKED (GLCNAC. . .)	(POTENTIAL).	186	197								
FT CARBOHYD 197	N-LINKED (GLCNAC. . .)	(POTENTIAL).	197	230								
FT CARBOHYD 230	N-LINKED (GLCNAC. . .)	(POTENTIAL).	230	234								
FT CARBOHYD 234	N-LINKED (GLCNAC. . .)	(POTENTIAL).	234	241								
FT CARBOHYD 241	N-LINKED (GLCNAC. . .)	(POTENTIAL).	241	262								
FT CARBOHYD 262	N-LINKED (GLCNAC. . .)	(POTENTIAL).	262	276								
FT CARBOHYD 276	N-LINKED (GLCNAC. . .)	(POTENTIAL).	276	289								
FT CARBOHYD 289	N-LINKED (GLCNAC. . .)	(POTENTIAL).	289	295								
FT CARBOHYD 295	N-LINKED (GLCNAC. . .)	(POTENTIAL).	295	301								
FT CARBOHYD 301	N-LINKED (GLCNAC. . .)	(POTENTIAL).	301	332								
FT CARBOHYD 332	N-LINKED (GLCNAC. . .)	(POTENTIAL).	332	339								
FT CARBOHYD 339	N-LINKED (GLCNAC. . .)	(POTENTIAL).	339	356								
FT CARBOHYD 356	N-LINKED (GLCNAC. . .)	(POTENTIAL).	356	386								
FT CARBOHYD 386	N-LINKED (GLCNAC. . .)	(POTENTIAL).	386	392								
FT CARBOHYD 392	N-LINKED (GLCNAC. . .)	(POTENTIAL).	392	397								
FT CARBOHYD 397	N-LINKED (GLCNAC. . .)	(POTENTIAL).	397	406								
FT CARBOHYD 406	N-LINKED (GLCNAC. . .)	(POTENTIAL).	406	448								
FT CARBOHYD 448	N-LINKED (GLCNAC. . .)	(POTENTIAL).	448	463								
FT CARBOHYD 463	N-LINKED (GLCNAC. . .)</											

RX MEDLINE=85228248; PubMed=2988795;
 RA Shaw G.M., Ganguly K., Gordon M., Conroy R., Schaber M., Kramers R.,
 RA Crowl R., Wong-Staal F., Reddy E.P.;
 RT "HTLV-III env gene products synthesized in E. coli are recognized by
 RT antibodies present in the sera of AIDS patients.";
 RL Cell 41:979-986(1985).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M14100; AAA44679.1; -.
 DR PDB: 1JAU; 17-OCT-01.
 DR PDB: 1JAV; 17-OCT-01.
 DR HIV: M14100; ENVSHXB3.
 DR InterPro: IPR000328; Env GP41.
 DR InterPro: IPR000777; GP120.
 DR Pfam: PF00516; GP120; 1.
 DR Pfam: PF00517; GP41; 1.
 DR AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
 KW 3D-structure.
 FT SIGNAL 1 30
 FT CHAIN 1 511
 FT CHAIN 512 856
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CARBOHYD 356 356
 FT CARBOHYD 386 386
 FT CARBOHYD 392 392
 FT CARBOHYD 397 397
 FT CARBOHYD 406 406
 FT CARBOHYD 448 448
 FT CARBOHYD 463 463
 FT CARBOHYD 611 611
 FT CARBOHYD 616 616
 FT CARBOHYD 624 624
 FT CARBOHYD 637 637
 FT CARBOHYD 674 674
 FT CARBOHYD 750 750
 FT CARBOHYD 816 816
 SQ SEQUENCE 856 AA; 97188 MM; 33730688B84C1AAC CRC64;
 Query Match 96.1%; Score 174; DB 1; Length 856;
 Best Local Similarity 97.3%; Pred. No. 2,46-16;
 Matches 36; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLRRAEAOQHLLQLTWQIKQIQANILAVERYLKQ 37
 Db 554 NLRRAEAOQHLLQLTWQIKQIQANILAVERYLKQ 590
 RESULT 8
 ID ENV_HVILW STANDARD; PRT; 856 AA.
 AC 070626;
 DT 15-JUN-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
 DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (HM12.3 isolate) (HIV-1).
 OC Viruses; Retroviruses; Retroviridae; Lentivirus.
 OC NCBI_TaxID=82834;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95127297; PubMed=7826699;
 RA Reitz M.S. Jr., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.M.,
 RA Shaw G.M., Kong L.I., Weiss S.H., Waters D., Gallo R.C., Blattner W.;
 RT "Viral variability and serum antibody response in a laboratory worker
 RT infected with HIV type 1 (HTLV type IIB)."
 RL AIDS Res. Hum. Retroviruses 10:1143-1155(1994).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U12055; AAA76690.1; -.
 DR PDB: 1IF3; 02-MAY-01.
 DR GLYCOsiteDB: 070626; -.
 DR InterPro: IPR000328; Env GP41.
 DR InterPro: IPR000777; GP120.
 DR Pfam: PF00516; GP120; 1.
 DR Pfam: PF00517; GP41; 1.
 DR AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
 KW 3D-structure.
 FT SIGNAL 1 30
 FT CHAIN 1 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT SIGNAL 1 30
 FT CHAIN 1 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT SIGNAL 1 30
 FT CHAIN 1 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT SIGNAL 1 30
 FT CHAIN 1 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT SIGNAL 1 30
 FT CHAIN 1 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT SIGNAL 1 30
 FT CHAIN 1 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT SIGNAL 1 30
 FT CHAIN 1 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT SIGNAL 1 30
 FT CHAIN 1 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT SIGNAL 1 30
 FT CHAIN 1 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT SIGNAL 1 30
 FT CHAIN 1 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT SIGNAL 1 30
 FT CHAIN 1 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT SIGNAL 1 30
 FT CHAIN 1 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT SIGNAL 1 30
 FT CHAIN 1 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT SIGNAL 1 30
 FT CHAIN 1 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT SIGNAL 1 30
 FT CHAIN 1 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT SIGNAL 1 30
 FT CHAIN 1 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT SIGNAL 1 30
 FT CHAIN 1 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT SIGNAL 1 30
 FT CHAIN 1 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT SIGNAL 1 30
 FT CHAIN 1 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT SIGNAL 1 30
 FT CHAIN 1 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT SIGNAL 1 30
 FT CHAIN 1 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBO

FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 856 AA; 96938 MW; 0C241332C7E6687 CRC64;

Query Match 96.1%; Score 174; DB 1; Length 856;
 Best Local Similarity 97.3%; Pred. No. 2,4e-16;
 Matches 36; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLRAIAEQHLLQLTWQIKQLQARILAVERYLKQ 37
 Db 554 NLRAIAEQHLLQLTWQIKQLQARILAVERYLKQ 590

RESULT 9
 ENV_HV1PV STANDARD; PRT; 856 AA.
 AC P03376;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Envelope polypeptide GP160 precursor [contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (PV22 isolate) (HIV-1).
 OC Viruses; Retrovirus; Retroviridae; Lentivirus.
 RX NCBI_Taxid=11700;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=8511157; PubMed=2982104;
 RA Muesing M.A., Smith D.H., Cabradilla C.D., Benton C.V., Laeky L.A., Capon D.J.;
 RT "Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovirus.";
 RL Nature 313:450-458(1985).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: K02083; AAB59873.1; -;
 DR EMBL: X01762; CA25903.1; ALT_SEQ.
 DR F03974; VCLJVL.
 DR HIV; K02083; ENVSPV2.
 DR InterPro: IPR000328; Env GP41.
 DR InterPro: IPR000777; GP120.
 DR Pfam: PF00516; GP120; 1.
 DR Pfam: PF00517; GP41; 1.
 DR AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal.
 KW SIGNAL.
 FT CHAIN 1 30 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT CHAIN 511 556 TRANSMEMBRANE GLYCOPROTEIN.
 FT DISULFID 54 74 BY SIMILARITY.
 FT DISULFID 119 205 BY SIMILARITY.
 FT DISULFID 126 196 BY SIMILARITY.
 FT DISULFID 131 157 BY SIMILARITY.
 FT DISULFID 218 247 BY SIMILARITY.

FT DISULFID 228 239 BY SIMILARITY.
 FT DISULFID 296 331 BY SIMILARITY.
 FT DISULFID 378 445 BY SIMILARITY.
 FT DISULFID 385 418 BY SIMILARITY.
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 856 AA; 97339 MW; 5FCDB1DC3C1209B3 CRC64;

Query Match 96.1%; Score 174; DB 1; Length 856;
 Best Local Similarity 97.3%; Pred. No. 2,4e-16;
 Matches 36; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLRAIAEQHLLQLTWQIKQLQARILAVERYLKQ 37
 Db 554 NLRAIAEQHLLQLTWQIKQLQARILAVERYLKQ 590

RESULT 10
 ENV_HV1BR STANDARD; PRT; 861 AA.
 AC P03377;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Envelope polypeptide GP160 precursor [contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (BRU isolate) (HIV-1).
 OC Viruses; Retrovirus; Retroviridae; Lentivirus.
 RX NCBI_Taxid=11686;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=85099333; PubMed=2981635;
 RA Wain-Hobson S., Sonigo P., Danos O., Cole S., Alizon M.;
 RT "Nucleotide sequence of the AIDS virus, LAV.";
 RL Cell 40:9-17(1985).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----

```
CC -----
DR EMBL; K02013; AAB59751.1; --
DR EMBL; A04321; CAA00352.1; --
DR PIR; A03975; VCLJLV.
DR PDB; 1ERE; 20-FEB-02.
DR HIV; K02013; ENVSEBU.
DR Interpro; IPR000328; Env_GP41.
DR Interpro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KM AIDS; Coat protein; glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 30
FT CHAIN 31 516
FT CHAIN 861 74
FT DISULFID 54 74
FT DISULFID 119 210
FT DISULFID 126 201
FT DISULFID 131 152
FT DISULFID 223 252
FT DISULFID 233 244
FT DISULFID 301 336
FT DISULFID 383 450
FT DISULFID 390 423
FT CARBOHYD 88 88
FT CARBOHYD 136 136
FT CARBOHYD 141 141
FT CARBOHYD 146 146
FT CARBOHYD 161 161
FT CARBOHYD 165 165
FT CARBOHYD 191 191
FT CARBOHYD 202 202
FT CARBOHYD 235 235
FT CARBOHYD 239 239
FT CARBOHYD 246 246
FT CARBOHYD 267 267
FT CARBOHYD 281 281
FT CARBOHYD 294 294
FT CARBOHYD 300 300
FT CARBOHYD 306 306
FT CARBOHYD 337 337
FT CARBOHYD 344 344
FT CARBOHYD 361 361
FT CARBOHYD 391 391
FT CARBOHYD 397 397
FT CARBOHYD 402 402
FT CARBOHYD 411 411
FT CARBOHYD 453 453
FT CARBOHYD 468 468
FT CARBOHYD 616 616
FT CARBOHYD 621 621
FT CARBOHYD 630 630
FT CARBOHYD 642 642
FT CARBOHYD 679 679
FT CARBOHYD 755 755
FT CARBOHYD 821 821
SQ SEQUENCE 861 AA; 97487 MM; 04DE2B4D4E6D5A CRC64;

Query Match 96.1%; Score 174; DB 1; Length 861;
Best Local Similarity 97.3%; Pred. No. 2,4e-16;
Matches 36; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxId=11691;
RN [1]
RP MEDLINE=90347835; PubMed=2384920;
RA Cheng-Mayer C, Quiroga M, Tung J W, Dina D, Levy J;
RT "Viral determinants of human immunodeficiency virus type 1 T-cell or
RT macrophage tropism, cytopathogenicity, and CD4 antigen modulation.";
RL J. Virol. 64:4390-4398 (1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M65024; AAA5072.1; --
DR PDB; 1OBE; 15-MAY-97.
DR HIV; M38428; ENVSEF162.
DR Interpro; IPR000328; Env_GP41.
DR Interpro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KM AIDS; Coat protein; glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 29
FT CHAIN 30 502
FT CHAIN 503 847
FT DISULFID 53 73
FT DISULFID 118 203
FT DISULFID 125 194
FT DISULFID 130 155
FT DISULFID 216 245
FT DISULFID 226 237
FT DISULFID 294 328
FT DISULFID 374 435
FT DISULFID 381 408
FT DISULFID 87 87
FT CARBOHYD 135 135
FT CARBOHYD 154 154
FT CARBOHYD 186 186
FT CARBOHYD 195 195
FT CARBOHYD 232 232
FT CARBOHYD 239 239
FT CARBOHYD 260 260
FT CARBOHYD 274 274
FT CARBOHYD 293 293
FT CARBOHYD 299 299
FT CARBOHYD 329 329
FT CARBOHYD 336 336
FT CARBOHYD 352 352
FT CARBOHYD 382 382
FT CARBOHYD 388 388
FT CARBOHYD 392 392
FT CARBOHYD 398 398
FT CARBOHYD 401 401
FT CARBOHYD 438 438
FT CARBOHYD 454 454
FT CARBOHYD 602 602
FT CARBOHYD 607 607
FT CARBOHYD 616 616
FT CARBOHYD 628 628
SQ SEQUENCE 847 AA; 96135 MM; 0A901317FD7EFA2B CRC64;

Query Match 95.6%; Score 173; DB 1; Length 847;
Best Local Similarity 94.6%; Pred. No. 3,3e-16;
```


Matches	35;	Conservative	1;	Mismatches	1;	Indels	0;	Gaps	0;
0y		1	NLRP1AENQOHILDTWQIKQOLQARILAVERYLKQD	37					
Db		545	NLRP1AENQOHILDTWQIKQOLQARILAVERYLKQD	581					
RESULT 12									
ENV_HV10Y		STANDARD;		PRT;		855	AA.		
AC	P20888;								
DT	01-FEB-1991	(Rel. 17, Created)							
DT	01-FEB-1991	(Rel. 17, Last sequence update)							
DE	15-JUL-1999	(Rel. 38, Last annotation update)							
DE	Envelope polyprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].								
GN	ENV.								
OS	Human immunodeficiency virus type 1 (OYI isolate) (HIV-1).								
OC	Viruses; Retroid viruses; Retroviridae; Lentiviruses.								
OX	NCBI_TaxID=11699;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RX	MEDLINE=90148544; PubMed=2559749;								
RA	Huet T, Dazza M.C., Brun-Vezinet F., Roelants G.E., Main-Hobson S.;								
RT	"A highly defective HIV-1 strain isolated from a healthy Gabonese individual presenting an atypical western blot.";								
RL	AIDS 3:707-715(1989).								
CC	-1- MISCELLANEOUS: THE OYI ISOLATE WAS TAKEN FROM THE BLOOD OF A HEALTHY GABONESE INDIVIDUAL.								
CC	-----								
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).								
CC	-----								
DR	EMBL; M26727; AAA83397.1; -.								
DR	HIV; M26727; ENVSOYI.								
DR	InterPro; IPR000328; Env GP41.								
DR	InterPro; IPR000777; GP120.								
DR	Pfam; PF00516; GP120.1.								
KW	AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal.								
FT	SIGNAL.	1	29						
FT	CHAIN	30	509						
FT	CHAIN	510	855						
FT	DISULFID	53	73						
FT	DISULFID	118	210						
FT	DISULFID	125	201						
FT	DISULFID	130	162						
FT	DISULFID	223	252						
FT	DISULFID	233	244						
FT	DISULFID	301	335						
FT	DISULFID	381	442						
FT	DISULFID	388	415						
FT	CARBOHYD	87							
FT	CARBOHYD	134	134						
FT	CARBOHYD	142	142						
FT	CARBOHYD	145	145						
FT	CARBOHYD	161	161						
FT	CARBOHYD	165	165						
FT	CARBOHYD	192	192						
FT	CARBOHYD	202	202						
FT	CARBOHYD	239	239						
FT	CARBOHYD	246	246						
FT	CARBOHYD	267	267						
FT	CARBOHYD	281	281						
FT	CARBOHYD	294	294						
FT	CARBOHYD	300	300						
FT	CARBOHYD	306	306						
EXTERIOR MEMBRANE GLYCOPROTEIN.									
TRANSMEMBRANE GLYCOPROTEIN.									
FT	DISULFID	53	73						
FT	DISULFID	118	210						
FT	DISULF								

Query Match	Best Local Similarity	94.6%	Score 173;	DB 1;	Length 855;
Matches 35;	Conservative	1;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1 NLRAIEAOOHLQLTWVOIKOLQARLIAREYHKDQ 37				
Dd	553 NLRAIEAOOHLQLTWVGIKOLQARLIAREYHKDQ 589				
RESULT 13					
ID	ENV_HV1JR	STANDARD;	PRT;	848 AA.	
AC	P208771;				
DT	01-FEB-1991 (Rel. 17, Created)				
DT	01-FEB-1991 (Rel. 17, Last sequence update)				
DT	10-OCT-2003 (Rel. 42, Last annotation update)				
DE	Envelope polypeptide GP160 precursor [containing: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP11)].				
DE	ENV.				
GN	Human immunodeficiency virus type 1 (JRCSF isolate) (HIV-1).				
OS	Viruses; Retroid viruses; Retroviridae; Lentivirus.				
OC	NCBI_TaxID=11688;				
RM	[1]				
RP	SEQUENCE FROM N.A. Koyanagi S., Chen I.S.Y.; Submitted (DEC-1988) to the HIV data bank.				
RA	Submitted (DEC-1988) to the HIV data bank.				
RL	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	-----				
CC	EMBL; M38429; AB03749.1; -. PDB; 1CE4; 18-MAR-99.				
DR	HIV; M38429; ENVJRCSF.				
DR	InterPro; IPR000328; Env GP41.				
DR	InterPro; IPR000777; GP120.				
DR	Pfam; PF00516; GP120.1.				
DR	Pfam; PF00517; GP41.1.				
KW	AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal; 3D-structure.				
FT	SIGNAL	1	32		
FT	CHAIN	33	503		EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	504	848		TRANSMEMBRANE GLYCOPROTEIN.
FT	DISULFID	53	73		BY SIMILARITY.
FT	DISULFID	118	203		BY SIMILARITY.
FT	DISULFID	125	194		BY SIMILARITY.
FT	DISULFID	130	154		BY SIMILARITY.
FT	DISULFID	216	245		BY SIMILARITY.
FT	DISULFID	226	237		BY SIMILARITY.
FT	DISULFID	294	328		BY SIMILARITY.
FT	DISULFID	374	437		BY SIMILARITY.
FT	DISULFID	381	410		BY SIMILARITY.
FT	CARBOHYD	87	87		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	134	134		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	137	137		N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 440 440 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 453 453 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 603 603 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 608 608 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 617 617 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 629 629 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 808 808 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 848 AA; 96475 MW; 20767F5127EC3F3 CRC64;

Query Match 94.5%; Score 171; DB 1; Length 848;
Best Local Similarity 91.9%; Pred. No. 6.2e-16;
Matches 34; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLRRAIEAOQHMLQITWQIKOLARVLAVERLYKQ 37
Db 546 NLRRAIEAOQHMLQITWQIKOLARVLAVERLYKQ 582

RESULT 14
ENV_HV1M STANDARD; PRT; 856 AA.
ID-ENV_HV1M
AC P05877;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (MN isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirinae.
OX NCBI_TaxId=11696;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88219542; PubMed=3369091;
RA Gargo C., Guo H.-G., Franchini G., Aldovini A., Collalti B.,
Farrell K., Wong-Staal F., Gallo R.C., Reitz M.S., Jr.;
"Envelope sequences of two new United States HIV-1 isolates.";
Virology 164:531-536(1988).
RT -1- MISCELLANEOUS: THE MN ISOLATE WAS TAKEN FROM A PEDIATRIC AIDS
PATIENT IN 1984.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; M17449; AAA44857.1; -
DR PDB; 1AC5; 31-JUL-94.
DR PDB; 1F58; 29-DEC-99.
DR PDB; 1NI2; 25-FEB-03.
DR PDB; 1NI2; 25-FEB-03.
DR HIV; M17449; ENV$MN.

```

```

DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
3D-structure.
FT SIGNAL 1 29
FT CHAIN 30 513
FT CHAIN 514 856
FT DISULFID 53 73
FT DISULFID 118 210
FT DISULFID 125 201
FT DISULFID 130 162
FT DISULFID 223 252
FT DISULFID 233 244
FT DISULFID 301 335
FT DISULFID 381 445
FT DISULFID 388 418
FT CARBOHYD 87 87
FT CARBOHYD 129 129
FT CARBOHYD 135 135
FT CARBOHYD 140 140
FT CARBOHYD 141 141
FT CARBOHYD 146 146
FT CARBOHYD 161 161
FT CARBOHYD 165 165
FT CARBOHYD 191 191
FT CARBOHYD 202 202
FT CARBOHYD 246 246
FT CARBOHYD 267 267
FT CARBOHYD 281 281
FT CARBOHYD 294 294
FT CARBOHYD 300 300
FT CARBOHYD 336 336
FT CARBOHYD 343 343
FT CARBOHYD 359 359
FT CARBOHYD 365 365
FT CARBOHYD 395 395
FT CARBOHYD 401 401
FT CARBOHYD 405 405
FT CARBOHYD 406 406
FT CARBOHYD 413 413
FT CARBOHYD 448 448
FT CARBOHYD 465 465
FT CARBOHYD 612 612
FT CARBOHYD 617 617
FT CARBOHYD 626 626
FT CARBOHYD 638 638
FT CARBOHYD 816 816
SQ SEQUENCE 856 AA; 97140 MW; D197D809940BE732 CRC64;

Query Match 94.5%; Score 171; DB 1; Length 856;
Best Local Similarity 91.9%; Pred. No. 6.2e-16;
Matches 34; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLRRAIEAOQHMLQITWQIKOLARVLAVERLYKQ 37
Db 555 NLRRAIEAOQHMLQITWQIKOLARVLAVERLYKQ 591

RESULT 15
ENV_HV12 STANDARD; PRT; 843 AA.
ID-ENV_HV12
AC P35961;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirinae.
OX NCBI_TaxId=36377;

```



```

FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 451 451 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 452 452 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 601 601 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 606 606 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 627 627 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 846 AA; 96476 MW; 8A3B9DA527DE2B83 CRC64;

Query Match 93.9%; Score 170; DB 1; Length 846;
Best Local Similarity 91.9%; Pred. No. 8.5e-16;
Matches 34; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLRAIEAOQHLLQLTWVGIKQLQARVLAVERLYKDQ 37
DB 544 NLRAIEAOQHLLQLTWVGIKQLQARVLAVERLYKDQ 580

RESULT 17
ENV_HV1B8 STANDARD; PRT; 851 AA.
AC P04582;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polypeptide GP160 precursor [contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP11)].
GN ENV.
OS Human immunodeficiency virus type 1 (BH8 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11684;
RN [1]
RX MEDLINE=8511123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,
RA Baumeister K., Ivanoff L., Petteway S.R., Pearson M.L.,
RA Lautenberger J.A., Papas T.S., Chrayeb J., Chang N.T., Gallo R.C.,
RA Wong-Staal F.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
RL Nature 313:277-284(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, K02011; AAA44661.1; .
DR PDB; 1DHR; 13-JAN-99.
DR PDB; 1HHG; 31-OCT-93.
DR PDB; 1QO3; 02-JAN-00.
DR PDB; 1S2T; 24-DEC-97.
DR HIV; K02011; ENV5B88.
DR GlycoSiteDB; P04582; .
DR InterPro; IPR00328; ENV_GP41.
DR InterPro; IPR00777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 30
FT CHAIN 1 506 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 507 851 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 54 74 BY SIMILARITY.
FT DISULFID 119 205 BY SIMILARITY.
FT DISULFID 126 196 BY SIMILARITY.
FT DISULFID 131 157 BY SIMILARITY.

```

```

FT DISULFID 218 247 BY SIMILARITY.
FT DISULFID 228 239 BY SIMILARITY.
FT DISULFID 296 331 BY SIMILARITY.
FT DISULFID 378 440 BY SIMILARITY.
FT DISULFID 385 413 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 443 443 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 606 606 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 632 632 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 669 669 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 745 745 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 811 811 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 851 AA; 96644 MW; D16A3C90857785F1 CRC64;

Query Match 93.9%; Score 170; DB 1; Length 851;
Best Local Similarity 94.6%; Pred. No. 8.6e-16;
Matches 35; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NLRAIEAOQHLLQLTWVGIKQLQARVLAVERLYKDQ 37
DB 549 NLRAIEAOQHLLQLTWVGIKQLQARVLAVERLYKDQ 585

RESULT 18
ENV_HV1S3 STANDARD; PRT; 852 AA.
AC P19549;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polypeptide GP160 precursor [contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (SF33 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11690;
RN [1]
RX MEDLINE=90317906; PubMed=2370688;
RA York-Higgins D., Cheng-Mayer C., Bauer D., Levy J.A., Dina D.;
RA "Human immunodeficiency virus type 1 cellular host range,
RT replication, and cytopathicity are linked to the envelope region of
RT the viral genome.";
RL J. Virol. 64:4016-4020(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```



```

FT DISULFID 130 152 BY SIMILARITY.
FT DISULFID 218 247 BY SIMILARITY.
FT DISULFID 228 239 BY SIMILARITY.
FT DISULFID 296 330 BY SIMILARITY.
FT DISULFID 376 444 BY SIMILARITY.
FT DISULFID 383 417 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 155 155 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 183 183 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 338 338 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 354 354 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 360 360 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 394 394 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 404 404 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 447 447 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 625 625 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 856 AA; 97526 MW; DB6BD1E49C404DB9 CRC64;

```

```

Query Match 93.9%; Score 170; DB 1; Length 856;
Best Local Similarity 91.9%; Pred. No. 8.6e-16;
Matches 34; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 NLRRAEAOHLLQLTWMOIKOLQARILAVERTLKQ 37
DB 554 NLRRAEAOHLLQLTWMOIKOLQARILAVERTLKQ 590

```

```

RESULT 22
ENV_HV1ZH STANDARD; PRT; 856 AA.
ID ENV_HV1ZH
AC P05881;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (Zaire H2321 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11692;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89228766; PubMed=2713163;
RA Strinivasan A., York D., Butler D. Jr., Jannoun-Nasr R., Getchell J.,
RT McCormick J., Ou C.Y., Myers G., Smith T., Chen E.;
RT "Molecular characterization of HIV-1 isolated from a serum collected
in 1976: nucleotide sequence comparison to recent isolates and
generation of hybrid HIV."
RT AIDS Res. Hum. Retroviruses 5:121-129 (1989).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```

CC EMBL; M15896; AAB53948.1; -
DR PIR; A44963; A44963.
DR HIV; M15896; ENV52321.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Signal;
KW Signal.
FT SIGNAL 1 29
FT CHAIN 30 511
FT CHAIN 512 856
FT DISULFID 53 73
FT DISULFID 118 206
FT DISULFID 125 197
FT DISULFID 130 153
FT DISULFID 219 248
FT DISULFID 229 240
FT DISULFID 297 331
FT DISULFID 379 445
FT DISULFID 386 418
FT CARBOHYD 87 87
FT CARBOHYD 132 132
FT CARBOHYD 138 138
FT CARBOHYD 152 152
FT CARBOHYD 156 156
FT CARBOHYD 183 183
FT CARBOHYD 198 198
FT CARBOHYD 242 242
FT CARBOHYD 263 263
FT CARBOHYD 277 277
FT CARBOHYD 294 294
FT CARBOHYD 302 302
FT CARBOHYD 332 332
FT CARBOHYD 339 339
FT CARBOHYD 355 355
FT CARBOHYD 364 364
FT CARBOHYD 387 387
FT CARBOHYD 393 393
FT CARBOHYD 398 398
FT CARBOHYD 402 402
FT CARBOHYD 411 411
FT CARBOHYD 448 448
FT CARBOHYD 461 461
FT CARBOHYD 462 462
FT CARBOHYD 465 465
FT CARBOHYD 611 611
FT CARBOHYD 616 616
FT CARBOHYD 637 637
SQ SEQUENCE 856 AA; 96909 MW; 8396E3F8BBD174E CRC64;

```

```

Query Match 93.9%; Score 170; DB 1; Length 856;
Best Local Similarity 94.6%; Pred. No. 8.6e-16;
Matches 35; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 NLRRAEAOHLLQLTWMOIKOLQARILAVERTLKQ 37
DB 554 NLRRAEAOHLLQLTWMOIKOLQARILAVERTLKQ 590

```

```

RESULT 23
ENV_HV1KB STANDARD; PRT; 861 AA.
ID ENV_HV1KB
AC P31819;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (KB-1 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.

```


FT CARBOHYD 254 254 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 275 275 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 344 344 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 351 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 367 367 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 407 407 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 413 413 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 455 455 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 471 471 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 634 634 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 646 646 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 825 825 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 865 AA; 97809 MW; 28828BC2314ADCAC CRC64;

Query Match 93.9%; Score 170; DB 1; Length 865;
 Best Local Similarity 91.9%; Pred. No. 8.7e-16;
 Matches 34; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLRRAEAOQHLLQLTWQIKOLARILAVERYLKQ 37
 Db 563 NLRRAEAOQHLLQLTWQIKOLARILAVERYLKQ 599

RESULT 25
 ENV_HV1J3 STANDARD; PRT; 867 AA.
 AC P12489;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Envelope glycoprotein GP160 precursor [contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_Taxid=11694;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=89352108; PubMed=2669897;
 RX Komiya N., Hattori N., Inoue J., Sakuma S., Kurimura T., Yoshida M.;
 RT "Nucleotide sequences of gag and env genes of a Japanese isolate of HIV-1 and their expression in bacteria."
 RL AIDS Res. Hum. Retroviruses 5:411-419(1989).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M21138; AB03526.1; -;
 DR HIV; M21138; ENV5JH3.
 DR InterPro; IPR000328; ENV_GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 KM AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
 KW signal.
 FT SIGNAL 1 30
 FT CHAIN 31 516 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT CHAIN 517 867 TRANSMEMBRANE GLYCOPROTEIN.
 FT DISULFID 53 73 BY SIMILARITY.

FT DISULFID 118 217 BY SIMILARITY.
 FT DISULFID 125 208 BY SIMILARITY.
 FT DISULFID 130 160 BY SIMILARITY.
 FT DISULFID 230 259 BY SIMILARITY.
 FT DISULFID 240 251 BY SIMILARITY.
 FT DISULFID 308 342 BY SIMILARITY.
 FT DISULFID 388 457 BY SIMILARITY.
 FT DISULFID 395 430 BY SIMILARITY.
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 159 159 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 253 253 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 366 366 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 408 408 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 412 412 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 423 423 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 475 475 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 622 622 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 627 627 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 636 636 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 648 648 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 867 AA; 98399 MW; 5F2310146B8E8680 CRC64;

Query Match 93.9%; Score 170; DB 1; Length 867;
 Best Local Similarity 94.6%; Pred. No. 8.7e-16;
 Matches 35; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NLRRAEAOQHLLQLTWQIKOLARILAVERYLKQ 37
 Db 565 NLRRAEAOQHLLQLTWQIKOLARILAVERYLKQ 601

RESULT 26
 ENV_HV1C4 STANDARD; PRT; 868 AA.
 AC P05879;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Envelope glycoprotein GP160 precursor [contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_Taxid=11687;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=87041461; PubMed=3490666;
 RX Desai S.M., Kalyanaraman V.S., Casey J.M., Srinivasan A.,
 RA Andersen P.R., Devare S.G.;
 RT "Molecular cloning and primary nucleotide sequence analysis of a distinct human immunodeficiency virus isolate reveal significant divergence in its genomic sequences."
 RT Proc. Natl. Acad. Sci. U.S.A. 83:8380-8384(1986).
 RL [2]
 RP SEQUENCE OF 34-43.

FT CARBOHYD 364 364 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 394 394 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 467 467 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 613 613 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 618 618 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 627 627 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 639 639 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 819 819 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 859 AA; 97109 MW; DBCF9AA52E3ABF29 CRC64;
 Query Match 93.4%; Score 169; DB 1; Length 859;
 Best Local Similarity 91.9%; Pred. No. 1.2e-15;
 Matches 34; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 NLRAIEAQOHLQLTWQIKQLQARILAVERYLKDQ 37
 Db 556 NLRAIEAQOHLQLTWQIKQLQARILAVERYLQDQ 592
 RESULT 28
 ENV HV1W2 STANDARD; PRT; 847 AA.
 AC POS880;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
 glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (WMJ2 isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 ON NCBI_TaxID=11705;
 RN [1]
 RP MEDLINE=86235450; PubMed=3012778;
 RA Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.R., Markham P.D.,
 RA Salanudin S.Z., Wong S.C., Gallo R.C., Parks E.S., Parks W.P.,
 RT "Genetic variation in HTLV-III/LAV over time in patients with AIDS or
 at risk for AIDS."
 RL Science 232:1548-1553 (1986).
 CC -1- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM
 CC BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO
 CC WAS PERINATALLY INFECTED BY HER MOTHER.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: M12507; AAB12990.1; -
 DR HIV; M12507; ENV\$WMJ2.
 DR InterPro: IPR000328; ENV_GP41.
 DR InterPro: IPR000777; GP120.
 DR Pfam: PF00516; GP120; 1.
 DR Pfam: PF00517; GP41; 1.
 KM AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
 KM Signal.
 FT CHAIN 1 29
 FT CHAIN 501 501 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT DISULFID 502 847 TRANSMEMBRANE GLYCOPROTEIN.
 FT DISULFID 53 73 BY SIMILARITY.
 FT DISULFID 118 202 BY SIMILARITY.
 FT DISULFID 125 193 BY SIMILARITY.
 FT DISULFID 130 152 BY SIMILARITY.
 FT DISULFID 215 244 BY SIMILARITY.
 FT DISULFID 225 236 BY SIMILARITY.

FT DISULFID 293 326 BY SIMILARITY.
 FT DISULFID 372 435 BY SIMILARITY.
 FT DISULFID 379 408 BY SIMILARITY.
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 155 155 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 231 231 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 273 273 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 334 334 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 380 380 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 450 450 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 602 602 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 628 628 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 847 AA; 96466 MW; CD1B33D73A5B0AE CRC64;
 Query Match 92.3%; Score 167; DB 1; Length 847;
 Best Local Similarity 89.2%; Pred. No. 2.2e-15;
 Matches 33; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 NLRAIEAQOHLQLTWQIKQLQARILAVERYLKDQ 37
 Db 545 NLRAIEAQOHLQLTWQIKQLQARILAVERYLQDQ 581
 RESULT 29
 ENV HV1Z8 STANDARD; PRT; 863 AA.
 AC POS882;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
 glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (Z-84 isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 ON NCBI_TaxID=11681;
 RN [1]
 RP MEDLINE=86281278; PubMed=3395517;
 RA Youniss J., Josephs S.F., Reitz M.S. Jr., Zagury D., Wong-Straal F.,
 RA Gallo R.C.,
 RT "Nucleotide sequence analysis of the env gene of a new Zairian
 RT isolate of HIV-1."
 RL AIDS Res. Hum. Retroviruses 4:165-173 (1988).
 CC -1- MISCELLANEOUS: THE Z-84 ISOLATE WAS TAKEN FROM A 54 YEAR-OLD
 CC ZAIREAN MALE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```

CC -----
DR EMBL; J03653; AAA44684.1; -.
DR HIV; J03653; ENV5JYL.
DR InterPro: IPR000328; Env GP41.
DR InterPro: IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polypotein; Glycoprotein; Transmembrane;
Signal.
FT SIGNAL. 1 29
FT CHAIN 30 518
FT CHAIN 519 863
FT DISULFID 53 73
FT DISULFID 118 217
FT DISULFID 125 208
FT DISULFID 130 160
FT DISULFID 230 259
FT DISULFID 240 251
FT DISULFID 308 342
FT DISULFID 388 452
FT DISULFID 395 425
FT CARBOHYD 87 129
FT CARBOHYD 129 136
FT CARBOHYD 136 142
FT CARBOHYD 142 143
FT CARBOHYD 143 159
FT CARBOHYD 159 163
FT CARBOHYD 163 194
FT CARBOHYD 194 199
FT CARBOHYD 209 209
FT CARBOHYD 209 246
FT CARBOHYD 246 274
FT CARBOHYD 274 288
FT CARBOHYD 288 301
FT CARBOHYD 301 307
FT CARBOHYD 307 343
FT CARBOHYD 343 350
FT CARBOHYD 350 365
FT CARBOHYD 365 396
FT CARBOHYD 396 402
FT CARBOHYD 402 406
FT CARBOHYD 406 412
FT CARBOHYD 412 455
FT CARBOHYD 455 468
FT CARBOHYD 468 469
FT CARBOHYD 469 472
FT CARBOHYD 472 618
FT CARBOHYD 618 623
FT CARBOHYD 623 632
FT CARBOHYD 632 644
FT CARBOHYD 644 823
FT CARBOHYD 823 823
SQ SEQUENCE 863 AA; 97743 MW; B729CB5A6FAD1641 CRC64;

Query Match 91.2%; Score 165; DB 1; Length 863;
Best Local Similarity 89.2%; Pred. No. 4.3e-15;
Matches 33; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 NLRLATEAOOHLQLTWYQIKOQLARILAVERYKQD 37
Db 561 NLRLATEAOOHLQLTWYQIKOQLARILAVERYKQD 597

```

```

OS Human immunodeficiency virus type 1 (1BR isolate) (HIV-1).
OC Viruses; Retrovirda; Retroviridae; Lentivirus.
OX NCBI_TaxID=11693;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=69085613; PubMed=2769516;
RA Andard R., Thayer R., Srinivasan A., Nayyar S., Gardner M., Luciw P.,
RT "Biological and molecular characterization of human immunodeficiency
RT virus (HIV-1BR) from the brain of a patient with progressive
RT dementia."
RT Virology 168:78-89(1989).
CC -1- MISCELLANEOUS: THIS VIRUS IS CYTOPATHICALLY ACTIVE AND WAS
CC HARVESTED FROM THE BRAIN TISSUE OF A NEUROLOGICAL AIDS PATIENT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M21098; AAA44221.1; -.
CC PIR; A31667; VCLBR.
CC PDB; 1IM7; 23-OCT-02.
CC HIV; M21098; ENV5JYLA.
CC InterPro: IPR000328; Env GP41.
CC InterPro: IPR000777; GP120.
CC Pfam; PF00516; GP120; 1.
CC Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polypotein; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL. 1 30
FT CHAIN 31 507
FT CHAIN 508 852
FT DISULFID 54 74
FT DISULFID 119 205
FT DISULFID 126 196
FT DISULFID 131 155
FT DISULFID 218 247
FT DISULFID 228 239
FT DISULFID 296 330
FT DISULFID 376 439
FT DISULFID 383 412
FT CARBOHYD 49 49
FT CARBOHYD 88 88
FT CARBOHYD 135 135
FT CARBOHYD 138 138
FT CARBOHYD 154 154
FT CARBOHYD 158 158
FT CARBOHYD 197 197
FT CARBOHYD 234 234
FT CARBOHYD 241 241
FT CARBOHYD 262 262
FT CARBOHYD 276 276
FT CARBOHYD 289 289
FT CARBOHYD 295 295
FT CARBOHYD 301 301
FT CARBOHYD 331 331
FT CARBOHYD 354 354
FT CARBOHYD 360 360
FT CARBOHYD 384 384
FT CARBOHYD 390 390
FT CARBOHYD 396 396
FT CARBOHYD 400 400
FT CARBOHYD 442 442
FT CARBOHYD 456 456
FT CARBOHYD 607 607
FT CARBOHYD 612 612
FT CARBOHYD 621 621
FT CARBOHYD 623 623
FT CARBOHYD 670 670
FT CARBOHYD 670 670

```

FT	CARBOHYD	812	812	N-LINKED (GLCNAC...)	(POTENTIAL)...
SQ	SEQUENCE	852 AA;	97203 MW;	2BB866345DCE915F	CRC64;

SQ SEQUENCE 852 AA; 97203 MW; 2BB866345DEC915F CRC64;

Query Match	89.5%	Score 162;	DB 1;	Length 852;
Best Local Similarity	86.5%	Pred NO. 1	1a-14.	

Best Local Similarity 86.5%; Pred.No. 1.1e-14;
Matches 32; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Matches 32; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Dy 1 NLRATEAQGHLLQLTWGIQIQLQARILAEVRYELKDQ 37
||| ||| : ||| ||| ||| : ||| |||
Db 550 NLMAIEAQGHMLLETWVGIGIQQLQAFVLAEVRYELKDQ 586

Db 550 NLLMAIEAQQHMLELTWGIKQIQARVLAVERYLKQ 586

Search completed: June 2, 2004, 11:42:57
Job time : 7.03261 secs

Job time : 7.03261 secs

THIS PAGE BLANK (USPTO)